QUINN EMANUEL URQUHART & SULLIVAN, LLP
Alex Spiro (admitted *pro hac vice*)
alexspiro@quinnemanuel.com
51 Madison Avenue, 22nd Floor
New York, New York 10010
Telephone: (212) 849-7000

QUINN EMANUEL URQUHART & SULLIVAN, LLP
Robert M. Schwartz (Bar No. 117166)
robertschwartz@quinnemanuel.com
Michael T. Lifrak (Bar No. 210846)
michaellifrak@quinnemanuel.com
Jeanine M. Zalduendo (Bar No. 243374)
jeaninezalduendo@quinnemanuel.com
865 South Figueroa Street, 10th Floor
Los Angeles, California 90017-2543
Telephone: (213) 443-3000

Attorneys for Defendant Elon Musk

UNITED STATES DISTRICT COURT

CENTRAL DISTRICT OF CALIFORNIA

VERNON UNSWORTH,

Plaintiff,

vs.

ELON MUSK,

Defendant.

Case No. 2:18-cv-08048

Judge: Hon. Stephen V. Wilson

DEFENDANT'S REPLY IN SUPPORT OF MOTION IN LIMINE NO. 5 TO EXCLUDE THE EXPERT OPINION OF DR. BERNARD J. JANSEN

Complaint Filed: September 17, 2018
Trial Date: December 2, 2019

Hearing Date: November 25, 2019
Time: 3:00 p.m.
Courtroom: 10A

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page

Cases

Rules and Regulations

I. INTRODUCTION

As the Plaintiff in a defamation case, seeking to justify a substantial award of damages, Mr. Unsworth wants to argue that a large number of people read Mr. Musk's statements about him. To that end, he wants Dr. Jim Jansen, supposedly a computer scientist, to tell the jury that this number was over 98 million. But Dr. Jansen cannot say that, because he does not know what the number is. Instead, he wants to testify to a different, misleading, and irrelevant number, based on methods that fail to satisfy *Daubert* and the Federal Rules of Evidence.

Plaintiff's opposition fails to rehabilitate the defects in Dr. Jansen's testimony. To begin, Plaintiff does not dispute that Dr. Jansen's "98 million" number does not represent the number of people who saw or read Mr. Musk's statements. Instead, it represents the number of people who merely ***visited*** websites that posted articles—***somewhere*** on the site—that mentioned those statements. Dr. Jansen has no idea ***where*** on those sites these articles appeared. Thus, he does not know if any of these articles were posted on the home page, or if someone had to scroll far down to find a link to them, or if they were not linked on the home page but were instead buried in the bowels of the site. That matters because websites contain hundreds and thousands of pages of content, and often many times more.

Plaintiff's opposition also does not dispute that ***Dr. Jansen does not know how many people actually saw or read any of these articles***, much less how many believed them or thought any less of Plaintiff. He knows he cannot say that number is 98 million. The information can be gathered. But Dr. Jansen chose not to find it.

Nothing in Plaintiff's opposition justifies wasting the jury's time with his irrelevant "site visitors" estimate, or risking that the jury will give it meaning it does not merit. Even if that were not the case, Plaintiff has failed to rebut the "junk science" foundations of Dr. Jansen's work. Plaintiff does not deny that, to estimate "site visitors," Dr. Jansen pulled the data from a website called SimilarWeb that estimates web traffic through an undisclosed and unknown sampling method that

has not been formally tested, has not been subject to peer review, has no known error rate, and that no court has deemed reliable. If Dr. Jansen had investigated these matters, as he should have, he would have be appalled by what he would have learned. When SimilarWeb's estimates have been scrutinized by independent researchers, SimilarWeb has been criticized for consistently and substantially overestimating traffic. And, apparently unbeknownst to Dr. Jansen, SimilarWeb also fails to take into account the enormous degree of web traffic—which he agrees is likely 30-40% of all web traffic—that does not result from human activity but is instead generated by bots. Obviously, the extent to which bots were "exposed" to Mr. Musk's statements about Plaintiff is of no concern to libel law.

Plaintiff's opposition also fails to meet his *Daubert* burden of establishing that each step in Dr. Jansen's methodology is scientifically reliable. Dr. Jansen's first step was to identify online articles that purportedly mention Mr. Musk's statements. He says he found 605. But that was not based on any scientifically reliable or testable methods. Instead, Dr. Jansen simply "conducted the investigation requested by Unsworth." (Opp. at 10.) And although Plaintiff's opposition attempts to dress up Dr. Jansen's methods to seem more scientific and rigorous than they are, no amount of fluff can change the fact that Dr. Jansen simply ran 5 to 12 Google searches, read the articles Google spit back, and divined their relevance, not as a computer scientist, but "as a thinking person." But each member of the jury is likewise a "thinking person" and should not be misled by his alchemy.

Plaintiff is also unable to defend Dr. Jansen's next step, namely, his use of an untestable and unreliable "method" to estimate web traffic from SimilarWeb data that fails three of the four *Daubert* factors. Plaintiff is unable to defend the fact that Dr. Jansen does not know what data SimilarWeb uses to estimate web traffic or how its algorithm then uses that data to generate its estimates.

The only evidence Plaintiff offers to support these methods undermines them. Dr. Jansen says he compared the SimilarWeb estimates to traffic results generated

by Comscore, an industry-leading service. But that document shows that SimilarWeb's traffic estimates for the same sites differed by ***34%***. Dr. Jansen is unable to account for this substantial discrepancy. The only other evidence he cites to support his use of SimilarWeb, from Screamingfrog, found that it overestimated traffic on 40% of the sites it analyzed, by an average of 17%. Yet Dr. Jansen identified no error rate and made no adjustment for these defects.

This is all the Court needs to hear to conclude that Dr. Jansen's method is not a reliable calculator of website "visits"—were that even relevant. The Court should bar Dr. Jansen from offering this unscientific testimony on an irrelevant topic.

II. DR. JANSEN'S TRAFFIC ESTIMATE IS NOT RELEVANT

To establish the extent to which his reputation was harmed, Mr. Unsworth will ask Dr. Jansen to testify that there were 98 million "daily unique visitors" in September 2018 to the websites that published articles containing Mr. Musk's statements about Mr. Unsworth. (Opp. at 8-9.) Dr. Jansen does not know how many people saw or read Mr. Musk's statements. (Dkt. No. 111-1, Jansen Depo. 30:24-33:13.) He does not even have ***actual*** daily visitor numbers for the sites he analyzed. His numbers are a summation of hundreds of third party (SimilarWeb) ***estimates***.

Plaintiff's opposition admits that Dr. Jansen "is not being offered to show how many people actually read Musk's statements." (Opp. at 8.) That should be enough to bar Dr. Jansen's testimony. Under California law, dissemination and republication require that the statement be "communicated to a third person who understands its defamatory meaning as applied to the plaintiff." *Tamkin v. CBS Broad., Inc.*, 193 Cal. App. 4th 133, 145-46 (2011).[1] Evidence of the availability of the statements is not evidence they were received and understood by anyone.

[1] Plaintiff argues that Defendant's motion mischaracterized *Tamkin*. Not so. It cited *Tamkin* for the rule that publication requires that the statement be communicated to a third party, who understands its meaning. That is what the case holds.

To avoid that result, Plaintiff cites a few cases for the proposition that a plaintiff does not need to plead that defamatory statements were actually read in order to establish the ***fact*** of publication. (Opp. at 7 citing *Farr v. Bramblett*, 132 Cal. App. 2d 36, 46-47 (1955); Opp. at 8.) But that is not the issue. Mr. Musk does not dispute that the statements were published. And Dr. Jansen is not being offered to establish the fact of publication. Mr. Unsworth wants his testimony to "prove[] damage and assist[] the trier of fact in determining the amount of damages to be awarded for the worldwide publication of the accusations." (Opp. at 1.) That is because he bears the burden to prove actual damages. *See Burnett v. Nat'l Enquirer, Inc.*, 144 Cal. App. 3d 991, 1013 (1983) (holding that, even in case of libel per se, subject to presumed damages, "[i]t remained nevertheless for respondent to establish the actual damage she had suffered as a result of the publication involved"). But Dr. Jansen's opinion—which does not establish how many people saw or read Defendant's statements, let alone formed a negative opinion of Plaintiff as a result—will not help the jury determine whether Plaintiff can meet that burden. (*See* Jansen Depo. 30:24-33:13.)

Plaintiff contends that he does not need to show that people saw or read about Mr. Musk's statements because the scope of "dissemination" can be used to support the award of damages. (Opp. at 9-10.) But none of the three cases he cites—each involving traditional print media or a TV news program—support such a result in this case. In the first place, none of those courts allowed the jury to hear testimony about ***estimated*** circulation or viewership numbers. In *Scott v. Times-Mirror Co.*, 181 Cal. 345, 365 (1919), *Weller v. Am. Broad. Companies, Inc.*, 232 Cal. App. 3d 991, 1013 (1991), and *Alioto v. Cowles Commc'ns, Inc.*, 430 F. Supp. 1363, 1371-72 (N.D. Cal. 1977), the plaintiffs identified the ***actual*** number of people who received the newspaper that published the defamatory remarks, the ***actual*** number of homes that the defamatory story was broadcast to, and the ***actual*** number of copies sold of the magazine containing the defamatory statement. That matters because "the best

place to get the amount of traffic that visited a site is from the website server itself." *McCurley v. Royal Seas Cruises, Inc*., 331 F.R.D. 142, 157 (S.D. Cal. 2019); *see also Finjan, Inc. v. ESET, LLC*, 2019 WL 5212394, at *4 (S.D. Cal. Oct. 16, 2019) (expert opinion based on projection was inadmissible as confusing and misleading where parties had access to actual numbers the projection purported to calculate).

That hardly begins to describe the lack of comparability between those cases and this one. In *Scott*, the defamatory statement appeared prominently on the front page of the *Los Angeles Times* City Sheet, and the newspaper was only 24 pages long. (*See* Supplemental Declaration of Michael Lifrak ("Supp. Lifrak Decl.") Ex. 8, *The Los Angeles Times,* Feb. 6, 1915, www.newspapers.com/image/380140241 (last visited Nov. 20, 2019).) In *Alioto*, the defamatory statement appeared on the cover of the magazine. (*See id*., Ex. 9, *Look Magazine,* Sept. 23, 1969, www.oldlifemagazines.com/look-magazine-september-23-1969-diana-ross.html (last visited Nov. 20, 2019.).) And in *Weller*, the defamatory statement was made in a TV news broadcast, which, given the nature of a TV broadcast, was passively received, in its entirety, by each household whose viewership was included in the Nielson number. In each case, it was not unreasonable for the court to allow the jury to consider the circulation or viewership numbers in deciding the extent to which the defamatory statement had been seen or heard.

But Dr. Jansen does not know whether anything comparable occurred here, that is, for example, if the articles in his collection were available on the "front" page of the websites whose visitor estimates he included, or how easy or hard it was for a typical visitor to any of those websites to come upon the Unsworth-related article he counted. He has no idea where any of these articles were located because he never bothered to look. (Jansen Depo. 146:2-18; 147:14-150:3;161:10-162:3.)

If he had, it would have collapsed his opinion. For example, on the day foxnews.com published a story about Mr. Musk's tweets, its home page included links to 92 separate pages, embedded 18 videos, and displayed headlines to more

than 150 articles, but did ***not*** include any mention or link to its story about Mr. Musk's tweets. (*See* Ex. 11, *July 15, 2018 Foxnews.com Home Page*, WAYBACK ARCHIVE, https://web.archive.org/web/20180715205857/http://www.foxnews.com/ (last visited Nov. 20, 2019).)

It is equally specious to treat the estimated traffic to a website as "a modern day equivalent of a newspaper's circulation." (Opp. at 9.) That is because a 24-page issue of the 1915 *Los Angeles Times* and a 1969 issue of *Look Magazine* bear no meaningful resemblance for this purpose to 2019-era websites. Modern websites are comprised of hundreds, thousands, or ***millions*** of pages of content, any individual page of which its visitors may navigate their way to in any number of different ways, most of which they will never see, and which sites more closely resemble encyclopedias or entire libraries than single issues of newspapers or magazines. (*See e.g.*, Supp. Lifrak Decl. Ex. 10, *New York Times Article Archive*, THE NEW YORK TIMES, www.nytimes.com/ref/membercenter/nytarchive.html (last visited Nov. 21, 2019) (stating that nytimes.com has more than 13 ***million*** online articles).

The foxnews.com website, discussed above, shows that someone who visited the home page to read the news of the day would not, like a reader of a newspaper or viewer of a TV news broadcast, have been exposed to the story about Mr. Musk's statements or a link to it. And according to Dr. Jansen's own documents, that is typically true for other websites because most visitors never go beyond the home page and, if they do, they do not go much further. (*See id*., Ex. 21 (SimilarWeb Website Performance Report for abcnews.com) (68% of visitors to abcnews.com do not go past the first page visited; average number of pages visited is 2.07); Ex. 22 (SimilarWeb Website Performance Report for cbsnews.com) (75% of visitors to cbsnews.com do not go past the first page visited; average number of pages visited is 2.38).)

Further, newspaper circulation numbers report the number of issues sold. It is not unreasonable to assume that, if someone is paying for a newspaper, the buyer will read it, even all of it. It is therefore not unreasonable to use its circulation number as some indication of the extent to which the defamatory statement was read. But many of the websites that Dr. Jansen included, such as abcnews.com, wikipedia.com, finance.yahoo.com, fortune.com, money.cnn.com, news.sky.com, nypost.com, people.com, slate.com, time.com, apnews.com, bbc.com, cnet.com, and cars.com, are available for free. Others are walled off to non-subscribers, or permit non-subscribers to see only a limited number of articles per month. In either case, each site contains thousands or millions of pages of content. Unlike a 24-page newspaper, it would be impossible for anyone to read them all or even more than a tiny percentage. Thus, it would be ***un***reasonable to use the number of visitors to the entire site as an indicator of the extent to which a particular article was seen or read.

To illustrate how improper and dangerous Dr. Jansen's methods and assumptions are, Defendant questioned him at his deposition on his inclusion of estimated visitor numbers to one of his sites, cars.com. Dr. Jansen deemed an article from that website to be relevant to his "dissemination" analysis, and it accounts for a large number of "visitors" in his total. (*See* Report, at 26.) Dr. Jansen admitted that the cars.com article he included was not found on the home page, or even linked from the home page. (Jansen Depo. 172:10-13.) And he acknowledged that the main reason people go to cars.com is to find a car to buy, not to read about what Elon Musk (or anyone else) had to say about Vernon Unsworth (or anyone else). (*Id.* at 164:15-166:25.) There are thousands of web pages on that site. Dr. Jansen has no idea how many people saw the article he included from cars.com. (*Id.* at 177:6-180:17; 30:24-33:13.) But the nature of that website and Dr. Jansen's admissions about why people visit it make it dangerous and prejudicial to allow the jury to rely on the estimated visitors to the entire cars.com website—

people who are there because they are shopping for a car—as evidence of anything in this case.

Whether for cars.com or any of the other websites whose overall user visit numbers Dr. Jansen includes in his 98 million figure, Plaintiff has done nothing to establish the necessary foundation for reliance on site data for the purpose he wants the jury to use it. And Dr. Jansen's testimony provides the jury with no way to process his traffic estimates, or to convert them into something the jury could reasonably rely on. Instead, jurors will have to fend for themselves in trying to evaluate these numbers. Plaintiff is counting on that, in the hopes that it will drive up the verdict. But this genuine risk of confusion and prejudice far outweighs any probative value of Dr. Jansen's "visitor" estimate and is why, when such concerns exist, courts exclude this kind of testimony. *See United States v. 87.98 Acres of Land in the City of Merced*, 530 F.3d 899, 906 (9th Cir. 2008) (excluding expert opinion as irrelevant and potentially confusing to the jury because it invited the jury to make inferences that were unsupported by the facts and the expert's testimony).

III. THE COURT SHOULD EXCLUDE DR. JANSEN'S OPINION THAT 605 ARTICLES CONTAINED DEFAMATORY STATEMENTS

Dr. Jansen identified 605 online articles that mention Mr. Musk's statements about Mr. Unsworth. As explained in the opening brief, this opinion is inadmissible under Rule 702 because Dr. Jansen failed to apply a reliable and testable methodology to reach this conclusion and because his "analysis" can be performed by a layperson. (Mot. at 12-15.) Plaintiff's defense of Dr. Jansen falls short.

First, Plaintiff dismisses Dr. Jansen's failure to apply reliable and testable methods to determine the number of articles by asserting that "Dr. Jansen conducted the investigation requested by Unsworth." (Opp. at 10.) That tells the Court nothing about the ***methods*** Dr. Jansen used. And it renders the method ***less*** reliable, not more. *See, e.g., United States v. Tran Trong Cuong*, 18 F.3d 1132, 1143 (4th Cir. 1994) ("Reports specifically prepared for purposes of litigation are not, by

definition, 'of a type reasonably relied upon by experts in the particular field.'") (quoting Fed. R. Evid. 703 (1975)).

Second, Plaintiff argues that Defendant "ignored" the steps that Dr. Jansen took to identify the articles and the "expertise" he applied. (Opp. at 11.) But the steps Plaintiff describes were the ones Dr. Jansen took to cast out a net to ***find*** articles to be included. (*Id.*) Plaintiff attempts to dress up those steps as "scientific" by asserting that he used "an algorithmic approach" to identify articles and continued to conduct Google searches until he reached "theoretical saturation," that is, until the Google searches were not turning up any new articles to consider including. (Opp. at 12.) That's just a fancy way of saying he ran a lot of Google searches, using the words Mr. Musk used to describe Mr. Unsworth, until he realized that Google was starting to return a lot of duplicates. (Jansen Depo. 204:20-207:17.) That takes no specialized training or expertise.

Third, and more important, Plaintiff's discussion (at pages 10-12 of the opposition) presents no defense of the steps Dr. Jansen ***then*** took to decide whether any of the articles should be ***included*** in his collection (and included in his tally of estimated visitors). The issue is not, as Plaintiff says, whether Dr. Jansen ***excluded*** articles that he should have included. (Opp. at 10-11.) The issue is the opposite: Dr. Jansen could not offer any basis—scientific or otherwise—for ***including*** the articles he included and, particularly, for how he decided that articles that discussed Mr. Unsworth's threats to sue or actual litigation did so only to "some" extent, but not so much that the articles were "primarily" about legal matters—a condition that, in his judgment, required him to not include the articles and not include the number of estimated visitors to the websites on which they were hosted. (*See* Report ¶ 58.) He admitted that those articles should be excluded. (Jansen Depo. 212:15-21.) Obviously, the articles were the result of Mr. Unsworth's actions, not Mr. Musk's. This issue infects a substantial part of his testimony. Of his 605 articles, 213 (or

35%) were published after Mr. Unsworth had either publically threatened to file this lawsuit or filed it. (*See* Supp. Lifrak Decl. Ex. 18.)

The *Daubert* problem for Dr. Jansen is that he created no "algorithm" and recorded no criteria that reflect the basis on which he decided to include the articles; he just read them and made up his mind. (Report ¶¶ 58-62; Jansen Depo. 52:4-53:10; 72:18-24; 223:8-224:9.) Although Plaintiff's opposition says he had criteria, he admitted at his deposition that he made his decisions, not based on his scientific expertise, but merely as a "thinking person." (Jansen Depo. 226:14-23.) He also admitted that he never memorialized any standards or criteria for including any of the 605 articles and created no record of the basis for any of the individual decisions he made to include any of them. (*Id.* at 223:8-224:9.) By failing to retain a log or notes on articles he picked—and those he rejected—he deprived Defendant and the jury of the ability to determine whether he reliably applied any meaningful criteria at all. (*See id.*) And absent some articulated and meaningful criteria for making those selections, which can be tested and replicated, his collection of articles (and the number of estimated visitors to the websites where they were hosted) is not the product of any reliable or scientific analysis. It's just a bunch of articles found by someone named Jim Jansen. This is inadmissible and unreliable lay testimony that does not satisfy *Daubert*. *See City of Pomona v. SQM N. Am. Corp.*, 750 F.3d 1036, 1047 (9th Cir. 2014) ("Under *Daubert's* testability factor, the primary requirement is that someone else using the same data and methods be able to replicate the results.") (internal quotes omitted); *see also Chesebrough-Pond's, Inc. v. Faberge, Inc.*, 666 F.2d 393, 398 (9th Cir. 1982) ("it would be well within the discretion of the trial judge" to exclude expert testimony regarding "a matter easily evaluated by laymen within the realm of their common knowledge and experience").[2]

[2] Plaintiff suggests that Dr. Jansen showed his expertise by pointing out at his deposition that the number of "results" Google returns do not necessarily mean

IV. THE COURT SHOULD EXCLUDE DR. JANSEN'S UNIQUE VISITORS ESTIMATE BECAUSE IT IS NOT RELIABLE

Dr. Jansen's method of estimating daily website visitors fail three of the four *Daubert* factors and is unreliable under Rule 702: (1) SimilarWeb has not been subject to any formal testing, (2) it has not been subject to peer review and publication, and (3) it has no known or calculated error rate. *See* Mot. at 9-12; *Daubert v. Merrell Dow Pharms., Inc.,* 509 U.S. 579, 593-594 (1993); *Allen v. Am. Capital Ltd.*, 287 F. Supp. 3d 763, 787-88 (D. Ariz. 2017) (excluding expert opinion that failed three of the four *Daubert* factors).

A. Dr. Jansen's Methods Are Not Reliable.

Plaintiff does not claim that any court has found SimilarWeb or Dr. Jansen's methods to be reliable under Rule 702.[3] He cannot, because no court ever has. Nor does Plaintiff claim that SimilarWeb's methods and algorithm have been tested by computer scientists to ensure their accuracy. There is no evidence that they have. Instead, Plaintiff asserts that some courts generally have found that, "where actual website data is not available, use of third party data collection programs can be reliable." (Opp. at 13.) So what? That does not mean that ***Dr. Jansen's*** particular data estimates here are reliable.

Moreover, the cases Plaintiff cites do not support this proposition. Instead, they hold that the use of ***actual*** web traffic data, not third party estimates or

---

"articles" and that the deeper one explores Google search results, the less likely each "result" links to a valid article. (Opp. at 11-12.) That information requires no expertise: it is found on the face of the Google search pages. Regardless, it has no bearing on any issue in the case.

[3] Plaintiff's opposition notes that Dr. Jansen "has been qualified to testify as an expert witness in one reported case." (Opp. at 3 citing *Wickfire, LLC v. Woodruff*, 2017 WL 1149075 (W.D. Tx. Mar. 23, 2017).) That is true. But in that case Dr. Jansen was not asked to estimate web traffic and he did not rely on SimilarWeb. *Id.* at *5. Instead he testified about IP addresses and clickable web ads. His qualification as an expert there has no bearing on his testimony here.

projections like SimilarWeb, is admissible to determine the number of people who visited a website. Thus, in *McCurley*, and contrary to Plaintiff's representation, the court "decline[d] to reach the issue of the reliability of converting Amazon's Alexa website traffic rankings into estimates of website traffic." 331 F.R.D. at 156. Instead, the court ***sustained*** the objection to the expert's website traffic ***estimates*** but allowed testimony based on "actual web traffic data from the servers of the websites themselves," which, unlike a third-party estimate, would be admissible. *Id.* at 157-158.

Similarly, in both *Schwartz v. Avis Rent a Car Sys., LLC*, 95 Fed. R. Evid. Serv. 350, at *4-6 (D.N.J. 2014), and *Granite State Trade Sch., LLC v. New Hampshire Sch. Of Mech. Trades, Inc.*, 120 F. Supp. 3d 56, 59-60 (D.N.H. 2015), the experts used the ***actual*** web traffic data and not some third-party estimate. Plaintiff and his expert could have obtained actual traffic data from the websites, but did not. The Court is not required to accept an unreliable alternative.

Plaintiff's reliance on *United States ex rel. Landis v. Tailwind Sports Corp.*, 2017 WL 5905509 (D.D.C. Nov. 28, 2017), is misplaced. The defendant in that case did not challenge whether the database from which the expert pulled his web traffic numbers used reliable methods to calculate visitors. *Id.* at *7. Thus, the court was not asked—and therefore did not consider—whether those methods for calculating web traffic had a known error rate, were subject to formal testing and peer review, or otherwise passed muster under *Daubert*. *Id.* The court did not consider whether the database included actual traffic numbers or estimates based on statistical sampling, like SimilarWeb, and it did not hold that third party statistical sampling is a reliable method for calculating web traffic. *Id.*

B. Dr. Jansen Is Unable to Establish SimilarWeb's Reliability.

Dr. Jansen has never worked for SimilarWeb. He does not know what data SimilarWeb collects or how that data is weighed to generate traffic estimates. He has never seen or analyzed SimilarWeb's algorithm. (Jansen Depo. 130:2-6;

131:25-132:15; 116:22-117:1; 118:1-18.) And Plaintiff concedes that Dr. Jansen has no "access to the proprietary database used by SimilarWeb" to create the web site traffic estimates he relied upon. (Opp. at 16.) Thus, Dr. Jansen is in no position to have "sufficiently described how it operates and why it is reliable." (Opp. at 16.)

Plaintiff asks the Court to look the other way. He argues that Dr. Jansen's ignorance of and lack of access to SimilarWeb's methods does not render his opinion defective. (Opp. at 17 citing *A&M Records, Inc. v. Napster, Inc.*, 2000 WL 1170106, at *6 (N.D. Cal. Aug. 10, 2000).) That is not the law. Courts regularly exclude expert opinions based on proprietary methods not subject to review or testing. *See*, *e.g.*, *Deficcio v. Winnebago Indus., Inc.*, 2014 WL 4211274, at *6 (D.N.J. Aug. 25, 2014) (excluding opinion based on proprietary method that "has not been subjected to peer review," has not been tested for accuracy, and has no known rate of error). And *Napster* is inapposite. There, the expert relying on proprietary software was the CEO of the company that created the software. *Napster*, 2000 WL 1170106, at *4. The other side could probe and test his underlying methods because, unlike Dr. Jansen, he had actual knowledge of and access to the software. *See id.* Here, however, Dr. Jansen can only refer to SimilarWeb's website and its general information. (Report ¶¶ 18 n.6, 24 n.9, 30 n.11-13, 74 n.48). Because he does not know how SimilarWeb arrives at its estimates and does not have access to that information, his method cannot be tested and cannot be admitted.

Plaintiff also contends that Dr. Jansen's methods "must" be reliable because SimilarWeb "has been relied on by experts and submitted by parties in reported cases." (Opp. at 17.) This is inaccurate and irrelevant. There is no record that SimilarWeb has ever been relied on by a qualified expert. The cases Plaintiff cites did not even involve expert testimony. *See In re Eros Int'l Secs. Litig.*, 2017 WL 6405846, at *5 (S.D. N.Y. Sept. 22, 2017) (SimilarWeb report attached to complaint); *Healthbox Global Partners, LLC v. Under Armour, Inc.*, 2016 WL

3919452, at *7 (D. Del. Jul. 19, 2016) (SimilarWeb report filed in opposition to motion for preliminary injunction). And neither of those courts credited the SimilarWeb report or found that it was reliable. In fact, in *Healthbox*, the court noted that SimilarWeb's report may have been incomplete and inaccurate. *Id.* at *7 n.15. Plaintiff thus can point to no case in which any court concluded that SimilarWeb is a reliable method of estimating web traffic.

As a fallback, Plaintiff cites four academic articles that mention or rely on SimilarWeb data. But these articles did not test or examine SimilarWeb's methods, data, or algorithm and therefore do not satisfy *Daubert*'s peer review prong. *Compare* Opp. at 18 *with Daubert*, 509 U.S. at 593 ("Another pertinent consideration is whether the theory or technique has been subjected to peer review and publication."). In fact, neither Plaintiff nor Dr. Jansen can identify any instance in which SimilarWeb's methods have been subject to peer review. (*See generally*, Report and Opp.) Two of the four articles merely cite SimilarWeb data in passing (*See* Supp. Lifrak Decl. Exs. 12-13) and the other two do not use SimilarWeb to estimate unique visitor traffic. (*See id*., Exs. 14-15.) Thus, they cannot support the reliability of Dr. Jansen's methods.

Plaintiff also argues that Dr. Jansen's opinion is reliable because he compared SimilarWeb's estimates to "actual data received from BuzzFeed" and "comparable data" from Comscore. (Opp. at 19.) But by Dr. Jansen's own admission, these two data points do not provide an apples-to-apples comparison to SimilarWeb's estimates. The BuzzFeed data does not reflect the number of daily unique visitors to BuzzFeed in September 2018. (Jansen Depo. 105:13-106:8.) And the Comscore estimate calculated all visitors to a collection of sites—for instance, instead of estimating traffic to foxnews.com, it estimates traffic to all Fox Corporation websites—and not a specific domain, as SimilarWeb attempts to do. (Report ¶ 76, n. 50.) The comparisons are therefore unreliable and cannot bolster the reliability of Dr. Jansen's methods. *See Superior Consulting Servs., Inc. v. Shaklee Corp.*, 2018

WL 2182303, at *3 (M.D. Fla. May 9, 2018) ("A comparison of web traffic from two different time periods that fails to match the domains being tracked during each time period is an unreliable one.").

Moreover, the fact that these two web traffic estimators generated vastly ***different*** visitor numbers for similar websites, as noted above, underscores the unreliability of these methods and the need for the Court to exercise its gatekeeping function here, particularly as his methods cannot be examined or tested.

C. Dr. Jansen Failed to Identify or Apply An Error Rate.

Dr. Jansen's method is unreliable and should be excluded because he failed to identify an error rate in SimilarWeb's analysis or apply one to its estimates. *See United States v. Cordoba*, 194 F.3d 1053, 1062 (9th Cir. 1999) (affirming exclusion of testimony regarding method with no known error rate); *United States v. Fultz*, 18 F. Supp. 3d 748, 758 (E.D. Va. 2014) (excluding report that was silent on the known or potential error rate in method); (*See generally*, Report; Jansen Depo. 201:25-202:6.) Plaintiff ignores this. The term "error rate" and the *Daubert*-required consideration of error rates appear nowhere in his brief.

Plaintiff fairs no better in addressing criticisms of SimilarWeb by the Ahrefs web metrics company and the failure of SimilarWeb and Dr. Jansen to account for bot traffic in their estimates. (Opp. 19-22.)

Ahrefs found that SimilarWeb's traffic estimation methods overestimated visitors on average by ***309***%. (*See* Dkt. No. 101-2, Lifrak Decl. Ex. 7 at 7.) Plaintiff has no answer. The method Ahrefs uses to determine SimilarWeb's accuracy is similar to the one used by the company known as Screamingfrog, which Dr. Jansen cites approvingly in his own report. (*Compare id*., Ex. 6 at 5-6 *with* Ex. 7 at 7.) Both compare SimilarWeb's estimates to actual traffic from approximately the same number of sites: Ahref looked at 24 sites and Screamingfrog looked at 25. (*Id.*) And both determined that SimilarWeb overestimated traffic. Screamingfrog found that SimilarWeb overestimated traffic for 40% of the sites, and that the net

overestimation, taking into account all traffic on all sites, was 17%. (*Id.*, Ex. 6.) Even if the Court ignored the Ahrefs overestimation of 309%, Dr. Jansen's own source confirms that SimilarWeb inflates its numbers. Even SimilarWeb itself acknowledges: "When it comes to online measurement, there is normally up to a 20% discrepancy between analytics tools," (Supp. Lifrak Decl. Ex. 17, *SimilarWeb vs. Direct Measurement*, SIMILARWEB, https://www.similarweb.com/blog/wp-content/uploads/2016/08/SW-vs-Direct-Measurement.pdf (last visited Nov. 21, 2019)), which makes it even harder to believe that Dr. Jansen adequately familiarized himself with the defects in the data on which he was relying and for which he chose not to identify or apply an error rate to account for this inflation. That fails *Daubert.*

Even worse, however, is Dr. Jansen's failure to account for non-human ("bot") traffic. He testified at his deposition that bots make up 30-40% of internet traffic. (Jansen Depo. 123:22-124:5.) Plaintiff argues that Defendant has mischaracterized Dr. Jansen's testimony. Not true. Dr. Jansen admitted that he did nothing to adjust SimilarWeb's numbers to account for bot traffic. (*Id*. at 134:3-18.) All he could say on the extent to which SimilarWeb's numbers (and, thus, his) were inflated by bot traffic, is that the "industry standard" is to exclude bot traffic from human traffic but that he could not state with certainty that SimilarWeb's estimates did so. (*Id.* at 121:18-123:3; 131:25-132:15.) In fact, Plaintiff is in no position to represent to the Court that SimilarWeb excludes bot traffic from its estimates. That is because SimilarWeb acknowledges, right on its website, that it does ***not*** attempt to exclude bot traffic from its estimates. (*See* Supp. Lifrak Decl. Ex. 16, *Similar Web vs. Google Analytics*, SIMILARWEB, https://support.similarweb.com/hc/en-us/articles/360001638917-SimilarWeb-vs-Google-Analytics (last accessed November 2, 2019) ("For example, some methodologies de-duplicate visits while SimilarWeb does not. Some tools discount or remove bot traffic, SimilarWeb does not.").)

Dr. Jansen should have investigated these matters before issuing his report and before testifying to subjects on which he claims to have expertise. Had he followed reliable and scientific methods, he would have known that he needed to apply an error rate to SimilarWeb's estimates to account for bot traffic. He failed to do so. Had he followed minimum standards in his field, he would have found out that SimilarWeb did not eliminate bot traffic, and thereby vastly overstated human traffic in his estimates, instead of attempting to bluff his way through cross-examination under oath. He failed to do that, too. He has not earned this Court's seal of approval in allowing him to present his opinions to the jury.

V. CONCLUSION

For the foregoing reasons, the Court should exclude Dr. Jansen's testimony.

DATED: November 21, 2019

Respectfully submitted,

QUINN EMANUEL URQUHART & SULLIVAN, LLP

By */s/ Alex Spiro*
Alex Spiro
Attorneys for Defendant Elon Musk